Amendment to the Specification:

Please amend the specification as follows:

Please replace the pending title with the following new title:

--PHOSPHATASE-ENCODING NUCLEIC ACIDS AND METHODS OF MAKING AND USING THEM--

Please replace the paragraph 36 on page 7, with the following amended paragraph:

A1

Figure 5 is an illustration of the full-length DNA (SEQ ID NO:19) and corresponding deduced [[deduded]] amino acid sequence (SEQ ID NO:28) of *Ammonifex degensii* KC4 of the present invention. Sequencing was performed using a 378 automated DNA sequence for all sequences of the present invention (Applied Biosystems, Inc. Foster City, California).

Please replace the paragraph 37 on page 7, with the following amended paragraph:

A2

Figure 6 is an illustration of the full-length DNA (SEQ ID NO:20) and corresponding deduced [[deduded]] amino acid sequence (SEQ ID NO:29) of *Methanococcus igneus* Ko15.

Please replace the paragraph 38 on page 7, with the following amended paragraph:

A3

Figure 7 is an illustration of the full-length DNA (SEQ ID NO:21) and corresponding deduced [[deduded]] amino acid sequence (SEQ ID NO:30) of *Thermococcus alcaliphilus* AEDII12RA.

Please replace the paragraph 39 on page 7, with the following amended paragraph:

A4 Figure 8 is an illustration of the full-length DNA (SEQ ID NO:22) and corresponding deduced [[deduded]] amino acid sequence (SEQ ID NO:31) of *Thermococcus celer*.

Please replace the paragraph 40 on page 7, with the following amended paragraph:

A5 Figure 9 is an illustration of the full-length DNA (SEQ ID NO:23) and corresponding deduced [[deduded]] amino acid sequence (SEQ ID NO:32) of *Thermococcus* GU5L5.

Please replace the paragraph 41 on page 8, with the following amended paragraph:

A6 Figure 10 is an illustration of the full-length DNA (SEQ ID NO:24) and corresponding deduced [[deduded]] amino acid sequence (SEQ ID NO:33) of OC9a.

Please replace the paragraph 42 on page 8, with the following amended paragraph:

A7 Figure 11 is an illustration of the full-length DNA (SEQ ID NO:25) and corresponding deduced [[deduded]] amino acid sequence (SEQ ID NO:34) of M11TL.

Please replace the paragraph 43 on page 8, with the following amended paragraph:

A8 Figure 12 is an illustration of the full-length DNA (SEQ ID NO:26) and corresponding deduced [[deduded]] amino acid sequence (SEQ ID NO:35) of *Thermococcus* CL-2.

Please replace the paragraph 44 on page 8, with the following amended paragraph:

A9 Figure 13 is an illustration of the full-length DNA (SEQ ID NO:27) and corresponding deduced [[deduded]] amino acid sequence (SEQ ID NO:36) of *Aquifex* VF-5.

Please replace the paragraph 244 on page 60, with the following amended paragraph:

A "comparison window", as used herein, includes reference to a segment of any one of the number of contiguous positions selected from the group consisting of from 20 to 600, usually about 50 to about 200, more usually about 100 to about 150 in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned. Methods of alignment of sequence for comparison are well-known in the art. Optimal alignment of sequences for comparison can be conducted, *e.g.*, by the local homology algorithm of Smith & Waterman, Adv. Appl. Math. 2:482, 1981, by the homology alignment algorithm of Needleman & Wunsch, J. Mol. Biol 48:443, 1970, by the search for similarity method of person & Lipman, Proc. Nat'l. Acad. Sci. USA 85:2444, 1988, by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by manual alignment and visual inspection. Other algorithms for determining homology or identity include, for example, in addition to a BLAST program (Basic Local Alignment Search Tool at the National Center for Biological Information), ALIGN, AMAS (Analysis of Multiply Aligned Sequences), AMPS (Protein Multiple Sequence Alignment), ASSET (Aligned Segment Statistical Evaluation Tool), BANDS, BESTSCOR, BIOSCAN (Biological Sequence Comparative Analysis Node), BLIMPS (BLocks IMProved Searcher), FASTA, Intervals & Points, BMB, CLUSTAL V, CLUSTAL W, CONSENSUS, LCONSENSUS, WCONSENSUS, Smith-Waterman algorithm, DARWIN, Las Vegas algorithm, FNAT (Forced Nucleotide Alignment Tool), Framealign, Framesearch, DYNAMIC, FILTER, FSAP (Fristensky Sequence Analysis Package), GAP (Global Alignment Program), GENAL, GIBBS, GenQuest, ISSC (Sensitive Sequence Comparison), LALIGN (Local Sequence Alignment), LCP (Local Content Program), MACAW (Multiple Alignment Construction & Analysis Workbench), MAP (Multiple Alignment Program), MBLKP, MBLKN, PIMA (Pattern-Induced Multi-sequence Alignment), SAGA (Sequence Alignment by Genetic Algorithm) and WHAT-IF. Such alignment programs can also be used to screen genome databases to identify polynucleotide sequences having substantially identical sequences. A number of genome databases are available, for example, a substantial portion of the human genome is available as

part of the Human Genome Sequencing Project ~~(J. Roach, http://weber.u.Washington.edu/~roach/human_genome_progress 2.html) (Gibbs, 1995)~~. At least twenty-one other genomes have already been sequenced, including, for example, *M. genitalium* (Fraser *et al.*, 1995), M. jannaschii (Bult *et al.*, 1996), *H. influenzae* (Fleischmann *et al.*, 1995), *E. coli* (Blattner *et al.*, 1997), and yeast (*S. cerevisiae*) (Mewes *et al.*, 1997), and *D. melanogaster* (Adams *et al.*, 2000). Significant progress has also been made in sequencing the genomes of model organism, such as mouse, *C. elegans*, and *Arabadopsis sp.* Several databases containing genomic information annotated with some functional information are maintained by different organization, and are accessible via the internet~~, for example, http://wwwtigr.org/tdb; http://www.genetics.wisc.edu; http://genome-www.stanford.edu/~ball; http://hiv-web.lanl.gov; http://www.ncbi.nlm.nih.gov; http://www.ebi.ac.uk; http://Pasteur.fr/other/biology; and http://www.genome.wi.mit.edu~~.

Please replace the paragraph 245 spanning pages 61 to 62, with the following amended paragraph:

One example of a useful algorithm is BLAST and BLAST 2.0 algorithms, which are described in Altschul *et al.*, Nuc. Acids Res. 25:3389-3402, 1977, and Altschul *et al.*, J. Mol. Biol. 215:403-410, 1990, respectively. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information ~~(http://www.ncbi.nlm.nih.gov/)~~. This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul *et al.*, *supra*). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always >0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero

or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, M=5, N=-4 and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength of 3, and expectations (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff & Henikoff, Proc. Natl. Acad. Sci. USA 89:10915, 1989) alignments (B) of 50, expectation (E) of 10, M=5, N= -4, and a comparison of both strands.

Please replace the paragraph 253 spanning page 63, with the following amended paragraph:

The BLAST programs identify homologous sequences by identifying similar segments, which are referred to herein as "high-scoring segment pairs," between a query amino or nucleic acid sequence and a test sequence which is preferably obtained from a protein or nucleic acid sequence database. High-scoring segment pairs are preferably identified (*i.e.*, aligned) by means of a scoring matrix, many of which are known in the art. Preferably, the scoring matrix used is the BLOSUM62 matrix (Gonnet *et al.*, Science 256:1443-1445, 1992; Henikoff and Henikoff, Proteins 17:49-61, 1993). Less preferably, the PAM or PAM250 matrices may also be used (see, *e.g.*, Schwartz and Dayhoff, eds., 1978, *Matrices for Detecting Distance Relationships: Atlas of Protein Sequence and Structure*, Washington: National Biomedical Research Foundation). BLAST programs are accessible through the U.S. National Library of Medicine~~, *e.g.*, at www.ncbi.nlm.nih.gov~~.

Please replace the paragraph 270 spanning page 68, with the following amended paragraph:

Figure 5 is a flow diagram illustrating one embodiment of an identifier process 300 for detecting the presence of a feature in a sequence. The process 300 begins at a start state 302 and then moves to a state 304 wherein a first sequence that is to be checked for features is stored to a memory 115 in the computer system 100. The process 300 then moves to a state 306 wherein a database of sequence features is opened. Such a database would include a list of each

feature's attributes along with the name of the feature. For example, a feature name could be "Initiation Codon" and the attribute would be "ATG". Another example would be the feature name "TAATAA Box" and the feature attribute would be "TAATAA". An example of such a database is produced by the University of Wisconsin Genetics Computer Group ~~(www.gcg.com)~~. Alternatively, the features may be structural polypeptide motifs such as alpha helices, beta sheets, or functional polypeptide motifs such as enzymatic active sites, helix-turn-helix motifs or other motifs known to those skilled in the art.

A13

Please replace page 73 (lacking a page number), with the attached page 73 (including the page number).

Please replace the paragraph 285 on page 73, with the following amended paragraph:

Ammonifex degensii KC4 - 3A1A

A14

5' CCGA GAA TTC ATT AAA GAG GAG AAA TTA ACT ATG GGG GCA GGT CCG AAA AGG 3' (SEQ ID NO:1)
5' CGGA GGA TCC CTA CAG TTC TAA AAG TCT TTT A 3'
Vector: pQET3 (SEQ ID NO:2)

Please replace the paragraph 286 on page 73, with the following amended paragraph:

Methanococcus igneus Ko15 - 9A1A

A15

5' CCGA GAA TTC ATT AAA GAG GAG AAA TTA ACT ATG TTG GAT ATA CTG [[GIT]] GTT 3' (SEQ ID NO:3)
5' CCGA CGA TCC TTA TTT TTT AAC CAA ATGT TCC 3'
Vector: pQET3 (SEQ ID NO:4)

Please replace the paragraph 287 on page 73, with the following amended paragraph:

Thermococcus Alcaliphilus AEDII12RA -18A

A16

5' CCGA CAA TTG ATT AAA GAG GAG AAA TTA ACT ATG ATG ATG GAA TTC ACT CGC 3' (SEQ ID NO:5)
5' CGGA GGA TCC CTA CAG TTC TAA AAG TCT [[TIT]] TTT A 3' (SEQ ID NO:6)
Vector: pQET3

Please replace the paragraph 288 on page 73, with the following amended paragraph:

Thermococcus Celer 25A1A (incorporating Mfe1 restriction site)

A17

5' CCGA CAA TTG ATT AAA GAG GAG AAA TTA ACT ATG AGA ACC CTG ACA ATA AAC 3' (SEQ ID NO:7)
5' CCGA GGA TCC TTA CAC CCA CAG AAC CCT TAC 3' (SEQ ID NO:8)
Vector pQET3

Please replace the paragraph 289 on page 73, with the following amended paragraph:

Thermococcus GU5L5 - 26AlA

A18

5' CCGA GAA TTC ATT AAA GAG GAG AAA TTA ACT ATG AAA GGA AAG TCT [[CIT GIT]] CTT GTT 3' (SEQ ID NO:9)
5' CCGA GGA TCC TCA AGC TTC CTG GAG AAT CAA 3' (SEQ ID NO:10)
Vector pQET3

Please replace the paragraph 290 on page 74 with the following amended paragraph:

OC9a - 27A3A

A19

5' CCGA GAA TTC ATT AAA GAG GAG AAA TTA ACT ATG CCA AGA AAT ATC GCC GCT 3' (SEQ ID NO:11)
5' CCGA GGA TCC TTA AGG [[CIT]] CTT CTC GAG GTG GGG [[GIT]] GTT 3' (SEQ ID NO:12)
Vector pQET3

Please replace the paragraph 291 on page 74 with the following amended paragraph:

M11 TL - 29A1A (incorporating Mfel restriction site)

A20

5' CCGA CAA TTG ATT AAA GAG GAG AAA TTA ACT ATG TAT AAA TGG ATT ATT GAG GG 3' (SEQ ID NO:13)
5' CCGA GGA CTA AAC ATA GTC TAA, GTA ATT AGC 3' (SEQ ID NO:14)
Vector pQET3

Please replace the paragraph 292 on page 74 with the following amended paragraph:

Thermococcus CL-2 30A1A

A21

5' CCGA GAA TTC ATT AAA GAG GAG AAA TTA ACT ATG AGA ATC CTC CTC ACC AAC 3' (SEQ ID NO:15)
5' CCGA GGA TCC TCA CAG GCT CAG AAG CCT TTG 3' (SEQ ID NO:16)
Vector pQET3

Please replace the paragraph 293 on page 74 with the following amended paragraph:

Aquifex VF-5 - 34A1A

A22

5' CCGA GAA TTC ATT AAA GAG GAG AAA TTA ACT ATG GAA AAC TTA AAA AAG TAC CT 3' (SEQ ID NO:17)
5' CCGA GGA TCC TCA CCG CCC CCT GCG GGT GCG 3' (SEQ ID NO:18)
Vector pQET3